

## Quick Guide for Users

Dataset:	Capital Stock
Dates available:	1980-2001
Source:	Businesses
Coverage:	All sectors
Collected by:	ONS
Link fields:	dlink_ref2 (IDBR)
Legal restrictions:	STA data

### Quick summary:

The ARD has no capital stock information. The suite of programs above use industry investment data and the ARD's own capital expenditure figures to create an estimate of capital stock for each reporting unit in each year 1980-2001. The estimates are calculated using the Permanent Inventory Method (PIM). Separate depreciation rates are used for plant and machinery, vehicles and buildings but an aggregate capital stock is reported.

The document *building the capital stock\_oct2002\_02.pdf* contains a description of the method, particularly in respect of the treatment of missing values, initial codes and firm births. It also shows the impact of some alternative assumptions.

The variables in the files are

year	year of observation
siclett	2-letter sic92
all_def95	average investment deflator (average over p&m., buildings and vehicles)
sel_idx	sel_idx=1 if capital stock could be computed for reporting unit
rcapstk95	estimate of real capital stock with 1995 base year
dlink_ref2	reporting unit identifier

### Sampling frame

The unit is the reporting unit from the ARD. In its current form, the data is available from 1980-2001.

### Organisation of files:

Files are created as *capstock\_xx.dta* in the *cap\_stock* folder of the data drive. The xx relates to the depreciation level on plant and machinery. As there is some dispute about the appropriate rate to use, various datasets have been created.

All years and sectors are aggregated into one file.

### Known data issues:

A change in the statistical treatment of leasing investment around 1988 occurred, but this seems not to have had any impact on the estimates.

Disinvestment is taken as evidence of prior investment when allocating initial shares. This may lead to a disproportionately high share of initial capital being allocated to declining or restructuring businesses.

### Other issues

Investigation of 1970's characteristics needs to be carried out.



## **Virtual Micro Data Laboratory**

# **Technical Guide: Estimating capital stock**

**Robert Gilhooly**

The Office for National Statistics (ONS) does not ask firms what their capital stocks are; hence, it is necessary to construct capital stock series and then link these to the other confidential data sets held within the Virtual Microdata Laboratory (VML). This technical guide is designed to be used in conjunction with the syntax file (.do) which created the capital stock data set. It provides a step-by-step guide to the key stages as set out in the syntax file, and follows the same ordering to aid understanding.

The Microdata Analysis & User Support (MAUS) team has recently made a new version of the Capital Stock data set available within the VML. It spans the years 1979-2005, and creates a firm-level estimate of capital stock for approximately 42,000 enterprises. This data set contains methodological improvements compared to previous versions, and, specifically, addresses the negative capital stock series which occurred in previous constructions of the data set.

The major development in the latest version of the capital stock data set is the ability for users to tailor the data set for their own purposes. The syntax file available within the VML has been designed to be very flexible; users can now easily create their own capital stock estimates by selecting from a range of key variables & methods, and re-running the file in their own work area. For example, by selecting a higher tolerance ratio of imputed to real investment figures, a user can significantly expand the number of enterprises for which a capital stock series are calculated.

This Technical Guide is intended to serve as a guide to the construction of the capital stock data set. It covers:

- The data sets used
- Dealing with missing values and imputation techniques
- Allocation of the aggregate capital stock series to firms' first appearance
- The Perpetual Inventory Model (PIM)
- Addressing negative capital stock series and injecting capital

The capital stock data set is unlike the majority of the other data sets held in the VML because it is a “constructed” data set. Therefore there are many decisions needed to overcome the issues outlined above, and there is not a definitive approach to take. The user of this data set is advised to read this document and then decide whether to use the file as it stands or adjust it using the options outlined.

## 1. Data sets

The capital stock data set is constructed from the Annual Respondents Database Panel Data (ARDPD) set, which is itself constructed from the Annual Respondents Database (ARD)<sup>1</sup>. The ARD holds firms' responses to the Annual Business Inquiry (ABI), while the MAUS team does some work on the raw ABI files to ensure that there is longitudinal consistency in variables and firm reference numbers.

The ABI is the most comprehensive business survey conducted by ONS and covers over 100 key economic variables, such as: turnover; production and operation costs; employment; industry classification, and investment. The ABI has surveyed all sectors since 1997, while firms in the manufacturing and construction sectors returned questionnaires in from 1973 and 1994 respectively. It is a census of large businesses, and a stratified sample of small and medium sized enterprises. The stratified sampling framework means that smaller firms move in and out of the survey, and to create capital stock estimates it is necessary to have a panel which accounts for such "missing years". The ARDPD solves this problem by creating entries for firms when they do not appear in the ARD, but are still trading e.g. if a firm is in the ARD in 1997 and 2000, the ARDPD ensures it is present for 1998 and 1999 also. This process creates a lot of missing values which must be imputed – in the previous example we would not have any investment figures for 1998 and 1999 to feed into the PIM – and section 2 outlines the approach we have taken to address this issue.

To create a PIM it is necessary to allocate shares of the aggregate capital stock series to firms for the first year that they appear in the survey (basically, the PIM depreciates this value, adds the net investment the following year and repeats these steps for all years). The aggregate capital stock series used is the Volume in Capital Services (VICS)<sup>2</sup>. Capital services are a flow measure which reflects the input of capital into production; thus, are more suitable for analysing productivity than the National Accounts wealth estimates of capital stock. The main difference between the National Accounts measure of capital stock and the capital services estimate is how the net stock estimates are weighted. Capital stocks are weighted using asset purchase prices while capital services are weighted using asset rental prices. The rental price is sometimes referred to as the user cost of capital, and it captures three components: depreciation of the asset, nominal rate of return and the change in the purchase price of the asset. If a firm is to rent an asset, the lease price must cover the loss in asset value due to depreciation; the more an asset depreciates the greater its rental price. Asset prices also affect the rental price as such movements reflect a capital gain or loss for the owner. Similar to previous, a declining asset price must be covered by a higher rental price to compensate the owner. Both of these key components can be illustrated by ICT assets. Prices of ICT goods continue to fall sharply, and a capital stock measure would result in such good being given less weight; however, this

---

<sup>1</sup> Detailed information about the data sets held in the VML can be found at:

<http://www.ons.gov.uk/about/who-we-are/our-services/vml/about-the-vml/datasets-available/dataset-downloads/index.html>

<sup>2</sup> More information about VICS can be found at:

<http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=14205&Pos=&ColRank=2&Rank=1000>

implies a rising rental price and a capital services measure would give such assets a higher weight. As noted by Wallis & Dey-Chowdhury (2007), there is a divergence between the volume of capital service and volume of capital stock series after 1980, which becomes increasingly stark after 1990. This divergence is caused by the move towards short-lived and more productive assets, such as computers, for which the flow of capital services is high. The standard capital stock measure does not capture this shift and so understates growth in the productive input of capital in the UK economy.

A real firm level capital stock series is constructed by MAUS using deflators to adjust for price changes in asset values over time. Deflators are constructed by SIC letter and asset type for each year, using 1995 as the base. The deflators are then applied to the investment figures (net capital expenditure) for each asset before they are used in the PIM. Real input and output series are required to remove price affects and reveal changes to firms' production efficiency.

## 2. Missing values and imputation techniques

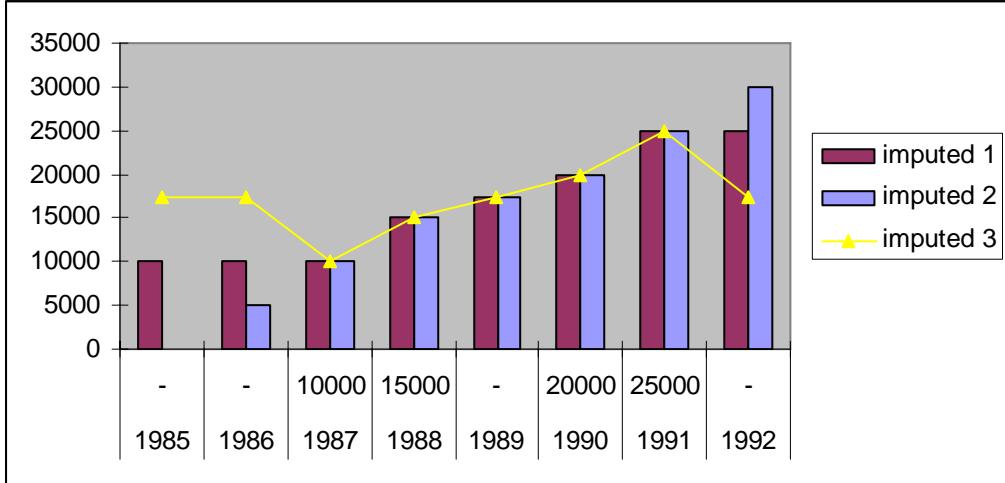
Some of the most important decisions to make in the construction of the capital stock data set are with regards to the treatment of missing cell values, and the cleaning of the data set before the PIM is run. There are many different approaches that we could take to this problem, and this section outlines how MAUS have chosen to impute cell values.

Employment is used as the spine for many of the imputations, plus the allocation of the aggregate capital share, thus the first stage is to ensure that we have a complete employment series. Generally, we can be confident using employment as this information is available from the Inter-Departmental Business Register (IDBR); hence, even if a firm does not respond to the questions in the ABI survey we are likely to have an employment figure. When a cell value is missing between years we simply interpolate the value. Missing (usually coded 0) cell values are more common in the first years of a company's existence, and these are filled in rolling back a three-year average of the employment figures available. This is illustrated in the table below:

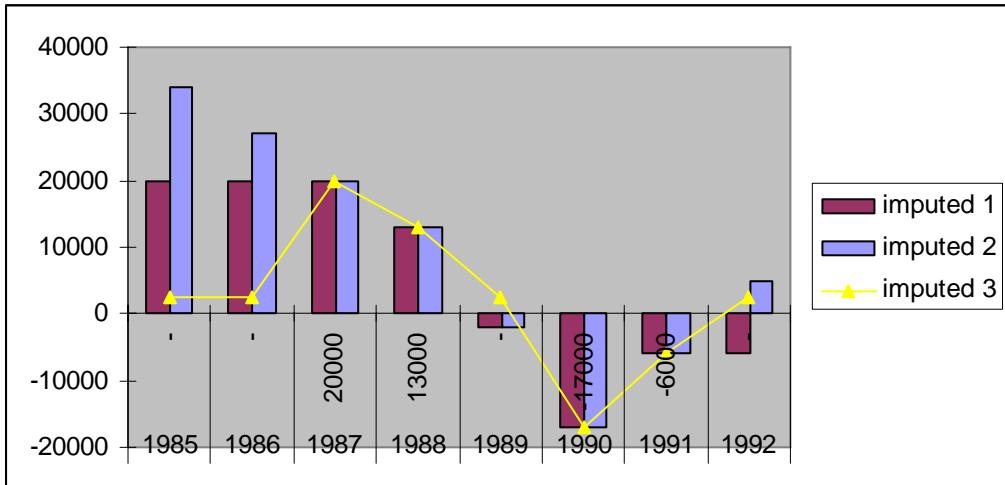
Year	Original cells		Final cells
	Employment	Adjusted Employment	
1980	0		108
1981	0		111
1982	0		113
1983	100		100
1984	120		120
1985	120		120
1986	114		114
1987	-		132
1988	150		150
1989	170		170
1990	165		165

We require a complete series of asset investment figures to compute the PIM and must also impute these missing values. The investment series in the ARD are net capital expenditure (capex) by asset. It is worth repeating that these figures are net i.e. acquisitions minus disposals. While we do ultimately require the net investment figure

for PIM, the volatile nature of firm investment and missing values provide a challenge for imputation – we are put further into the dark as acquisition and disposals series, which might provide more information on underlying investment patterns, are not available. The following graphs illustrate the problems we face:



The graph above shows 3 methods of imputing missing values (labeled “-“; while the figures on the x-axis represent real values e.g. 15,000 for 1988) for a firm which has four positive net capex cells from 1985 to 1992. Imputation 1 interpolates the 1989 value and provides a constant extrapolation at the edges (1985, 1986 & 1992)<sup>3</sup>. Imputation 2 also interpolates, but provides momentum extrapolation at the edges. Imputation 3 creates an average capex figure over the real values and then uses this number to fill-in the missing cells. At this stage, the three capex series already paint different pictures, which will impact on the estimated capital stock series produced by the PIM. It is critical to stress the importance of the net position i.e. negative values are likely.<sup>4</sup> The graph below shows what happens to the imputed series when net capex becomes negative in some years:



The capex values given above are typical of the values we see in the microdata, and reflect the volatile nature of firm investment patterns. The firm shown above is a net investor (+10k) over the four years for which we have data; it is not clear which

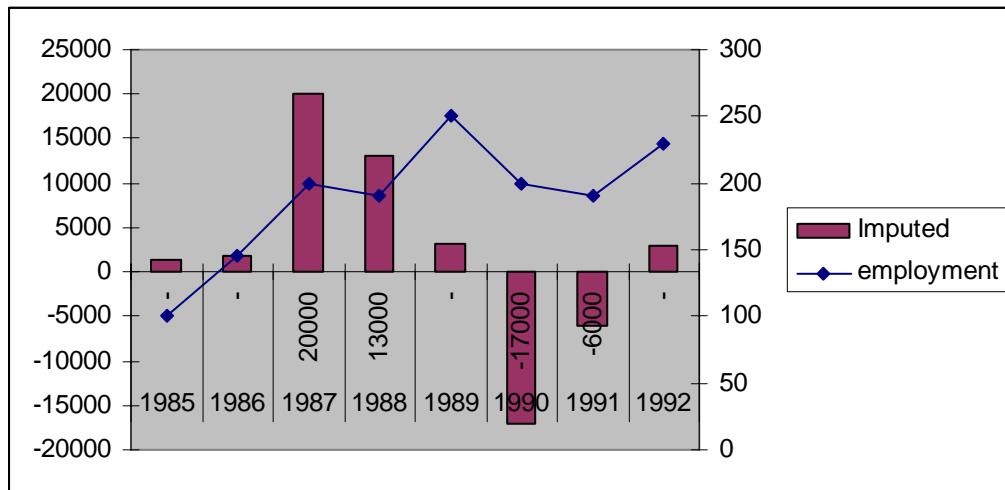
<sup>3</sup> This technique is used in Martin (2003).

<sup>4</sup> The use of net investment partially counters criticism of constructing capital stock estimates at the Reporting Unit level i.e. firms which close local units should report a negative investment. However, as illustrated by Harris (2005), firms will not be allocated additional capital when they acquire new local units. This is another reason why the capital stock series may underestimate the true figures.

imputation technique (if any) could be deemed to be “best” given the volatile nature of firm investment. Methods 1 and 2 give strong investment figures, but the true values could equally be negative. It should be noted that (almost) all imputation techniques will smooth the estimated investment patterns below the volatile true series which we would see if we had the real investment figures for the missing years.

The MAUS team decided to use the employment series to help with the imputation of capex. It is hoped that this gives the estimated capex figures more real information compared to just using one of the techniques outlined above. The latest capital stock data set thus creates a capex series which inputs using an average of net capital expenditure per employee for the real data, and then applies this to the missing values. This is similar to imputation method 3 above, and while it produces low investment figures for missing years it does ensure some consistency in the net position i.e. net (positive) investors continue to invest over the years. This technique is increasingly valuable in highly volatile series, where the other methods often impute large negative figures. The table and graph below illustrate the imputation for a firm which is a net investor for the four years which we have real data.

year	capex	employment	Imputed
1985	-	100	1282
1986	-	145	1859
1987	20000	200	20000
1988	13000	190	13000
1989	-	250	3205
1990	-17000	200	-17000
1991	-6000	190	-6000
1992	-	230	2949
<b>Averages</b>	<b>2500</b>	<b>195</b>	<b>n/a</b>
<b>Average capex per employee</b>			<b>13</b>



While it is possible to impute for any number of missing values the quality of the imputation will suffer when we have fewer real values to base our calculations on (i.e. we are likely to impute disposals of capital for all years for some firms). In practice, it is prudent to set a tolerance level. Currently the maximum ratio of imputed to real values is set at a maximum of 1:1. Lowering our intolerance to imputed values will

create estimates of firm capital stock for many more firms, and is likely to pick-up more SMEs due to the stratified sampling framework within the ABI. The table below outlines the number of firms for which we calculate capital stock given different tolerance levels. Users can easily alter the “calc\_ratio” in the syntax file and re-run to suit their needs.

Calc_ratio	Tolerance level (max number of imputed to real values)	Number of firms with capital stock series
0.33	1:3	13131
0.66	2:3	21809
1	1:1	41782
1.5	3:2	50622
2	2:1	74356

### 3. Allocating capital to firms

The first stage in the PIM is to calculate the share of total industry capital which should be allocated to the firms in your sample. We do this by assuming that the ratio of firms’ investment (i.e. sum of individual firms’ investment) to industry investment corresponds to their share of industry capital. The calculation of the initial capital allocation is not precise, and allocating firms too little capital is one reason for the implausible generation of negative capital stock series (see section 5). It seems plausible that the stratified sampling framework of the ABI/ARD results in too low an allocation of industry capital. The largest firms – presumably with most capital – represent a disproportionate number of the firms in the microdata. The formulae for the first stage of capital stock allocation is given by:

$$\text{Allocation firm Capital}_{(asset)} = \text{Industry Capital}_{(asset)} * \text{Firm Investment Share}_{(asset)}$$

$$\text{Firm Investment Share}_{(asset)} = \frac{\sum (rncapex)_{year, siclett, asset}}{\text{Industry Investment}_{year, siclett, asset}}$$

Now that we have decided how much capital should be allocated to the firms we are interested in, the second step is to devise a method for splitting this capital amongst the firms. At this junction we want to choose a variable which is plausibly related to capital stock and is available for all the years. The main options are “total purchases” and “materials & fuels” – both could be seen as proxies for the capital stock and firms are asked these questions in all years. Alternatives, which are only available from 1997 onwards, include: “number of local units”, “Spending on Insurance” and “Spending on Road Transport”. As noted above, the new flexible .do file allows users to specify which variable they would like to use, and within this .do file there is another .do file which compares the alternative options against the number of capital stock series which are created and are negative. The capital stock data set created by the MAUS team indicates that the best results, for all years, are obtained by specifying the key variable to be “total purchases”. A basic allocation of firm level capital can be created using:

$$m0\_Share = \left[ \frac{\text{key var}}{\sum \text{key var}_{siclett, year}} \right]$$

We shall refer to the basic formulae above as Method 0 (m0) in the syntax file. Method 0 is only able to split capital at the level of SIC letters, which is another reason why capital stock series may become negative. Consider firms in the same SIC letter but different SIC codes (e.g. a1=Civil engineering, a2=plumbing):

Method	Firm #	Year	SIC letter	SIC code	total purchases	Share of capital
m0	1	1998	a	a1	20	0.11
m0	2	1998	a	a1	20	0.11
m0	3	1998	a	a2	50	0.26
m0	4	1998	a	a2	100	0.53

We can make an attempt to (partially) remedy this situation by using a more advanced method of allocating capital to firms. We can weight the allocation of capital by both Total Purchases and Employment (or an alternative variable). The formulae below outline two alternative methods for creating a share measure:

$$m1\_Share = \left[ \frac{totpurch}{\sum_{sic3d, year} totpurch} \right] * \left[ \frac{\sum employment_{sic3d, year}}{\sum employment_{siclett, year}} \right]$$

$$m2\_Share = \left[ \frac{employment}{\sum employment_{sic3d, year}} \right] * \left[ \frac{\sum totpurch_{sic3d, year}}{\sum totpurch_{siclett, year}} \right]$$

The tables below illustrate how the 2 different methods create different shares compared to the original methodology.

Method	Firm #	SIC letter	SIC code	Total purchases	Employment	SIC code: TP	SIC code by letter: Employment	Share of capital
m1	1	a	a1	20	50	0.50	0.10	0.05
m1	2	a	a1	20	25	0.50	0.10	0.05
m1	3	a	a2	50	200	0.33	0.90	0.30
m1	4	a	a2	100	500	0.67	0.90	0.60

Method	Firm #	SIC letter	SIC code	Total purchases	Employment	SIC code: Employ.	SIC code by letter: Total purchases	Share of capital
m2	1	a	a1	20	50	0.67	0.21	0.14
m2	2	a	a1	20	25	0.33	0.21	0.07
m2	3	a	a2	50	200	0.29	0.79	0.23
m2	4	a	a2	100	500	0.71	0.79	0.56

The flexible set-up of the capital stock syntax allows a user to specify which method (m0, m1 or m2) they would like to use in their capital stock. Investigation by the MAUS team indicates that Method 1 produces the best results. See table below:

Counts of the number of negative capital stock series created under different methods				
method	All assets	Plant & Machinery	Vehicles	Buildings
m0_totpurch	229	173	1719	207
m0_matfuel	622	444	4888	943
m1_totpurch	216	164	1652	189
m1_matfuel	613	446	4719	853
m2_totpurch	293	220	2596	333
m2_matfuel	393	282	4525	689

This table illustrates one of the major problems in constructing the capital stock data set. The volatility in the net investment figures results in over 10% of the capital stock series for vehicles becoming negative. Obviously, it is impossible for a firm to hold a negative stock of vehicles, and this points to a combination of the main problems: not allocating enough initial capital; an inability to split capital at a detailed enough level; imputation techniques failing to create a large enough investment to assets (e.g. vehicles) when there are missing cell values; and, possibly, using too high a depreciation rate.

In the current methodology, all firms are allocated a share of the aggregate capital stock when they first enter the data set. The PIM which we specify also adds on the firm's investment which it undertakes in its first year. It could be argued that if an investment figure is available we should just use that as the firm's first year capital stock; however, it seems highly probable that the firm was in existence before it is picked up by the ABI survey, thus allocating a share of the aggregate capital stock figure in its first year provides a proxy for a firm's "pre-ABI" capital stock.

#### 4. Perpetual Inventory Model

After selecting your key variable (total purchases), methodology (capital share allocations) and dealing with missing values, running the PIM is relatively uncontentious. All that is left for the user to do is to select the depreciation rates they would like to use. The MAUS team follow the depreciation rates set out by ONS: Plant & Machinery 6%, Buildings 2%, Vehicles 20%. However, alternatives such as those based on BEA estimates include: Plant & Machinery 13%, Buildings 2.5%, Vehicles 25%.<sup>5</sup>

The PIM uses geometric depreciation as follows:

$$\text{Firm Capital Stock}_{asset,year_i} = (\text{Allocation firm Capital}_{asset} * \text{Firm Investment Share}_{asset}) + \text{rncapex}_{year_i}$$

$$\text{Firm Capital Stock}_{asset,year_{i+1}} = [\text{Firm Capital Stock}_{asset,year_i} * (1 - \text{asset depreciation})] + \text{rncapex}_{year_{i+1}}$$

$$\text{Firm Total Capital Stock}_{year} = \sum \text{Firm capital}_{asset}$$

We choose to use geometric depreciation, because the asset price depreciation and physical deterioration (age-efficiency profile) of the assets coincide.<sup>6</sup>

<sup>5</sup> See Fraumeni (1997)

<sup>6</sup> See ONS Productivity Handbook (2007)

## 5. Dealing with the remaining negative capital stock series

Given that a negative capital stock series is impossible in reality, we can address this problem by injecting additional capital at a firm level. This can be done by stipulating that a firm's capital stock must not go below zero, increasing the investment figures in the proceeding years and then re-running the PIM. It is advisable to inject the capital close to the year in which the capital stock turns negative; consider the huge capital injection which would be required if a 20 year vehicles series turned negative in the last year i.e. a -£1000 figure would need £87,000 ( $1000 * [1/0.8]^{20}$ ) injection in the first year to ensure that an entirely positive series is created. For this reason we localize the injection to at most three years previously:

$$\text{Asset Investment Injection} = (|\text{negative capital stock}|) * \left[ \frac{1}{(1 - \text{depreciation})} \right]^n$$

The table below illustrates how a firm's vehicle stock may turn negative – assuming 20% depreciation – and how a localized injection of capital can ensure that stocks do not fall below zero:

year	net capex_vehicles	Capital allocation	capital stock	adjusted capex	adjusted capital stock
1980	10,000	50,000	60,000	10,000	60,000
1981	12,000	0	60,000	12,000	60,000
1982	-5,000	0	43,000	-5,000	43,000
1983	-2,000	0	32,400	-2,000	32,400
1984	20,000	0	45,920	<b>52,986</b>	78,906
1985	40,000	0	76,736	40,000	103,125
1986	-20,000	0	41,389	-20,000	62,500
1987	-50,000	0	<b>-16,889</b>	-50,000	<b>0</b>
1988	20,000	0	6,489	20,000	20,000
1989	5,000	0	10,191	5,000	21,000
1990	5,000	0	13,153	5,000	21,800

In practice, the calculated capital stock series often becomes negative for many observations within a series; hence, it is often necessary to run the above injection several times. The current syntax file calls in the “localised\_boost.do” file several times in a row to deal with this problem.

## **Capital stock** **Methodological and data Issues**

- 1) Deflators for “poth”, “pcom” and “int” are aggregated with a simple mean. A weighted mean should be more appropriate (We could use VICS investment as weights)
- 2) Small reporting units are sampled on a random basis hence the ARD displays missing values in those years in which small firms are not selected. In order to avoid underestimating the accumulation of capital by small firms, before applying the PIM, missing investment values are linearly interpolated.

This methodology implies that firms follow a smooth capital adjustment pattern over time. Despite this assumption is in line with neoclassical investment models, there is evidence that firms adjust their capital stocks in a lumpy manner. Investment seems to occur in large and concentrated episodes. (Doms and Dunne, 1998 and Attanasio et Al, 2003). Worth investigating investment patterns to see if assumption behind interpolations are correct? Worth considering alternative methods (ie regressing missing investment)?

- 3) The sectoral capital stock produced for the construction of 2003 Volume Index of Capital Services (Vaze, 2003) separates computer from plant and machinery and apply the appropriate deflators and depreciation rates. Keeping computers separate is quite important because: i) ICT assets have a shorter life-span than other type of assets; ii) they are characterised by rapid fall in relative price due to quality improvements (see Oulton and Srinivasan, 2003 and Tevlin and Whelan, 2000).

When constructing the capital stock at reporting unit level we loose this level of detail. In fact, since in the reporting unit investment in plant&machinery includes computer software and hardware, when entering VICS data, we aggregate the sectoral information for plant and machinery (poth) with hardware (pcom) and software (int) in order to make it consistent with the ARD information. As a consequence, when running the PIM the capital stock for plant and machinery is incorrectly depreciated and deflated. Worth investigating with the ABI team the possibility of constructing separate investment data for computer?

- 4) Depreciation rates should be updated using the information available in VICS.xls. An average depreciation rate for “poth”, “pcom” and “int” should be calculated taking into account the relative importance of each asset. Furthermore, depreciation rates should be kept with a split by asset and sector (current version is split by assets only)

5) When constructing the initial capital stock at enterprise-level, the sectoral capital stock is split among all reporting unit in sector j and year 1. Then, the investment is added to the initial capital stock. This causes a double counting problem because the investment in the initial year is already accounted for in the sectoral capital stock. One way to solve this problem would be to lag the sectoral investment and relative shares by one period:

$$K_{iajt} = K_{aj(t-1)} * \varphi_{aj(t-1)} * \gamma_{ij(t-1)} + I_{iajt}$$

t=age=1

$\varphi$ =share of sectoral capital stock to be assigned to selected sample

$\gamma$ =share of selected sample capital stock to be assigned to the individual reporting unit

6) At the end of the programme, the capital stock by asset is aggregated to obtain an overall measure of capital stock: “rcapstk95”. When aggregating different assets we should use appropriate weights (currently we just sum up the 3 measures of capital stock by asset). According to Oulton and Srinivasan (2003), each asset should be weighted according to its asset’s price.

## Bibliography

Attanasio, O. P., L. Pacelli, I. Reduto dos Reis (2003) “Investment Patterns in UK manufacturing Establishments”,

Doms, M. and T. Dunne (1998) “Capital Adjustment Patterns in Manufacturing Plants, Review of Economic Dynamics”, n.1, 409-429

Oulton, N. and S. Srinivasan (2003) “Capital stocks, capital services, and depreciation: an integrated framework”, Bank of England, Working Paper n.192

Tevlin, S. and K. Whelan (1999) “Explaining the Equipment Investment Boom of the 1990s”, Division of Research and Statistics, Federal Reserve Board, December 29, mimeo.

Vaze, P. (2003) “Estimates of the volume of capital services”, Economic Trends, n.600

## Construction of Capital stock

### Description of programme

The construction of the capital stock involves the use of the following do-files:

- 1.prepare\_deflators.do
- 2.make\_row.do
- 3.pim.do

#### 1. **prepare\_deflators.do**

- Input **deflators** from **VICS.xls**; these are implied deflators from historic investment series calculated by the ONS for the construction of 2003 Volume Index of Capital Services (VICS). The Volume Index of Capital Service measure the flow of capital services derived from all the capital assets that exist in a sector (see Vaze, 2003).

Deflators are split by asset, industrial sector and year:

##### Assets:

**b** = Buildings

**poth** = Plant other (plant and machinery including purchased software and excluding hardware)

**pcom** = Computer hardware

**v** = Vehicles

**int** = Intangibles (dominated by own-account software, excluding purchased software)

##### Industrial sectors:

The industrial desaggregation is identical to the Supply Use Table (SUT)'s classification for Gross Fixed Capital Formation with a breakdown of 36 industries (see information in **VICS.xls** for details)

- “poth”, “pcom” and “int” are averaged to create one asset type called “pm”, representing plant and machinery assets. In fact, we need to aggregate the 3 assets in order to make VICS information consistent with the enterprise-level information from the ARD.
- If deflator is missing it is replaced with the average deflators across sectors by year;
- Drop observations <1979
- Finally create one file per asset type (each split by sector and year):

**b\_def95.dta** (deflator for buildings)

**pm\_def95.dta** (deflator for plant and machinery)

**v\_def95.dta** (deflator for vehicles)

## 2. make\_row.do

Input sectoral [net capital stock](#) from VICS.xls. This is the capital stock net of geometric depreciation estimated by the ONS for the construction of 2003 VICS. Net Capital Stock of an asset is a wealth measure calculated summing up the past history of gross investment in that asset in constant prices and subtracting estimated depreciation and retirements. Depreciation is what must be spent to maintain the value of capital stock at the existing level.

- “poth”, “pcom” and “int” are aggregated into one asset type called “pm”.
- drop observation <1979
- transform capital stock’s unit from millions to thousands pound to make it consistent with ARD investment data (which are in thousands)
- assign 0 if capital stock is negative (this is to avoid negative shares)
- saved as “[prabstock02.dta](#)”

Siclett	Year	ind_capstk95b	ind_capstk95pm	ind_capstk95v
---------	------	---------------	----------------	---------------

- Input sectoral [investment](#) in constant price from VICS.xls. This is the Gross Fixed Capital Formation (GFCF) in 1995 prices used by the ONS for the construction of 2003 VICS. The GFCF consists of resident producers' acquisitions less disposals on fixed capital assets (as above split by asset type, industrial sector and year)
  - “poth”, “pcom” and “int” are aggregated into one asset type called “pm”.
  - drop values<1979
  - If invest is = 0 it is replaced with a missing value(zero values would not allow us to estimate the sectoral share of capital stock to be assigned to the sample of selected reporting units later on in the programme)
  - Missing values are linearly interpolated
  - Take absolute values of investment (this is to avoid negative shares)
  - Transform investment’s unit from millions to thousands pound
  - Save as [prabinvest.dta](#)

Siclett	Year	ind_invest95b	ind_invest 95pm	Ind_invest 95v
---------	------	---------------	-----------------	----------------

- Input reporting unit-level information from the [ARD Standard Variables](#):

dlink_ref2	reporting unit reference number (key consistent over time based on post 94 idbr_ref)
totpurch	total purchases of goods and services
ncapex	total net capital expenditure*
ncapex_pm	net capital expenditure* for plant and machinery
ncapex_b	net capital expenditure* for land & buildings
ncapex_v	net capital expenditure* for vehicles
go	Gross output (used for checks only)
sel_emp	employment figures held to sampled enterprises (used for checks only)

\*In this context “net capital expenditure” means acquisitions less disposal proceeds of fixed assets. It excludes any allowances for depreciation. See “code&derivation” file for more details on capex variables.

- `ncapex_pm` is not reliable; thus, it is replaced with a residual figure calculated from "total net capital expenditures" as follows:

**`ncapex_pm=ncapex-(ncapex_b+ncapex_v)`**

"`ncapex`", derived from total acquisition (q600 ABI) less disposal, includes land, buildings, plant, machinery (hence computer hardware), purchased computer software, and computer software developed by own staff (see ABI survey form for more detail). Consequently, "`ncapex_pm`" represents an estimate for investment in plant and machinery, including hardware and software.

- Save as `rpanel_cap.dta`
- Load the Register Panel dataset (`reg_pan.dta`). The Register Panel, which contains selected and non selected reporting unit all in one file, allows us to identify and keep reporting units which are always selected and those which are selected only in some years (hence drop units which are never selected). In years in which the small enterprises aren't selected we will interpolate the missing investment values.
- Merge Register Panel variables with Standard Variables (merge `rpanel_cap.dta` using `dlink_ref2`)

Keep following variables:

```
dlink_ref2
year
ncapex (renamed ncapex_all)
ncapex_pm
ncapex_b
ncapex_v
sel_id (marker which indicates if reporting unit was selected in a specific year)
sel_emp
go
totpurch
sic92
region
```

- Correct wrong asset types in 1994 & 1998 by assigning missing values.
- Create age variable =`_n` (this variable will be used to identify when the reporting unit enters the dataset for the first time. This information is required when estimating the initial capital stock at reporting unit level)
- Create SIC92 with 2 digit from SIC92-4 digit
- SIC 92-2 digit are converted into SUT classifications using "numlettlook\_7" (See VICS.xls for info). This conversion is necessary to make ARD industrial classification consistent with VICS industrial classification.
- Input deflators
  - Create an average deflator across all assets (`all_def95`)
  - Use deflators to create real investment at reporting unit level:

```
rncapex_all95=ncape_all/all_def95  
rncapex_b95= ncape_b/b_def95  
rncapex_pm95= ncape_pm/pm_def95  
rncapex_v95= ncape_v/v_def95
```

- Interpolate investment missing data (when missing at the edge take previous or first non missing) and generate “rnkapex\_aa95\_old” to distinguish original data from the interpolated one “rnkapex\_aa95” (aa=asset)
- In order to identify reliable interpolation and original values, generate sel\_idx==1 if the maximum number of interpolations per reporting unit are less than 6 and investment values (*ncapex\_all95*) are non-missing. This marker will be used at the end of the programme to restrict the calculation of capital stocks to “reliable” investment figures only.
- Input investment and capital stock data at sector level **prabstock02.dta**
- Calculate the initial share of sectoral capital stock to be assigned to the sample of selected units (from now on we will refer “selected” as selected by the survey and interpolated). In fact, since our sample of units represent only a certain percentage of enterprises in the sector, we need to calculate how much of the sectoral capital stock we need to allocate to our sample of units in each sector (and year). We calculate this share using the following hypothesis:

Capital share of selected sample = Investment share of selected sample =  $\varphi$

Calculate the investment share of selected sample in every sector/year as a percentage of sectoral investment in corresponding sector/year :

$$\varphi_{ajt} = \text{invest\_inflator}'aa'_{jt} = \frac{\sum |I_{iajt}|}{|I_{ajt}|}$$

- Calculate reporting unit's share of initial capital stock:

This share is estimated assuming the following hypothesis:

Reporting unit's initial capital share = Reporting unit's average material share.

- Calculate material share:

$P_i$  = reporting unit purchase of goods and services (totpurch)

- By year and sector, creates the sum of total purchases across all units:

$$\text{sum\_totpurch}_{ijt} = \sum_{i \in jt} P_{(i)} \quad \text{if sel\_idx==1}$$

- Reporting unit material share:

$$M_{ijt} = \frac{P_{jti}}{\sum_{i \in jt} P_{ijt}}$$

- Calculate reporting unit average material share (average across life span of the unit)

$$\bar{M}_i = \frac{\sum M_{ijt}}{T}$$

- Calculate reporting unit average material share relative to year- totals.

$$\gamma_{ijt} = \frac{\bar{M}_i}{\sum_{i \in jt} M_i} = \frac{\frac{1}{T} \sum_{t \in i} M_{ijt}}{\sum_{i \in jt} \frac{1}{T} \sum_{t \in i} M_{ijt}}$$

save as **cappanel\_raw**

### 3.pim.do

- Set local depreciation rates:  
Local pm=0.06
- Set depreciation rates by asset:

$$\begin{aligned} d\_all &= (pm+b+v)/3 \\ d\_pm &= 0.06 \\ d\_b &= 0.02 \\ d\_v &= 0.2 \end{aligned}$$

- Generate empty capital stock at reporting unit level-> rcapstk\_aa'95 -> K<sub>ijt</sub>
- Estimate total capital stock to be assigned to the selected sample in each year/sector. This is estimated multiplying VICS capital stock by the investment share of selected unit in corresponding year/sector

rem\_`aa'= ind\_capstck95 aa' \* invest\_inflator aa'

rem\_aa' =  $K_{ajt} * \varphi_{ajt}$  , make it equal to 0 if negative

- Estimate initial capital stock at reporting unit level. This is obtained multiplying the capital stock of all selected units in sector j and year t=age=1 (year in which the unit first appears in the sample) by the average material share of the individual reporting unit, then adding the reporting unit's investment in same year

if t=age=1 & sel\_idx=1

rcapstk\_aa'95= rem\_aa'\*mat\_sh\_av\_sh+rncapex\_aa'95

$$K_{iaj1} = K_{aj1} * \varphi_{aj1} * \gamma_{ij1} + I_{iaj1}$$

- Apply Perpetual Inventory Method for all years: capital stock in previous year less capital consumption (depreciation) plus investment in current year

If sel\_idx=1

$$K_{iajt} = K_{iaj(t-1)} * (1 - d_a) + I_{iajt}$$

rcapstk\_pm95 = net capital stock for plant and machinery (including comp. hardware & software)

rcapstk\_b95= net capital stock for land and buildings

rcapstk\_v95= net capital stock for vehicles

rcapstk95=rcapstk\_pm95+rcapstk\_b95+rcapstk\_v95

save capstock\_pm\_dep'